A Bayesian approach to estimating detection performance in a multi-sensor environment

Steven A. Horn
Antony Zegers
DRDC Centre for Operational Research and Analysis

D.W. Mason
Edward J. Emond
DRDC Centre for Operational Research and Analysis Emeritus

Defence Research and Development Canada

Scientific Report
DRDC-RDDC-2014-R56
September 2014

A Bayesian approach to estimating detection performance in a multi-sensor environment

Steven A. Horn
Antony Zegers
DRDC Centre for Operational Research and Analysis

D.W. Mason
Edward J. Emond
DRDC Centre for Operational Research and Analysis Emeritus

Defence Research and Development Canada

Scientific Report
DRDC-RDDC-2014-R56
September 2014

Abstract

This report presents the results of Operational Research support to maritime surveillance operations in the Canadian maritime approaches. The principal development is a Bayesian method to estimate the performance of sensors in a way that can be applied during ongoing operations. This novel method does not require knowledge of the sea-truth to evaluate sensors. A method to estimate sea-truth is also presented, which provides a new and unique analysis capability for the estimation of surveillance gaps. The series of new metrics and methods are implemented in a variety of operational tools for analysis of the recognized maritime picture (RMP), which support new operational procedures and processes. The desired result is a more robust and optimized maritime surveillance capability.

Significance to defence and security

This report presents a body of research conducted under Defence Research & Development Canada (DRDC) Advanced Research Project (ARP) 11hn aimed at improving maritime surface vessel surveillance. The methods developed take advantage of having multiple sensors surveying a given region at a given time period to produce estimates of sensor detection probabilities and estimates of the number of targets in the region that might have gone undetected.

A major side-benefit of the Bayesian approach when applied to sensor mix analysis is that it can also be used to estimate the (generally unknown) sea-truth. A closed form solution has been derived for the distribution of the total number of targets that may have been out there in any given multi-sensor trial (the union of all detections plus the undetected targets). This solution is in the form of the Negative Binomial distribution. This provides a useful estimate of the performance of any given collection of sensors employed to feed the recognized maritime picture (RMP).

Impact

This research has led the Canadian Forces to adopt a policy of employing multiple sensors simultaneously in surveillance exercises and trials on the West Coast in order to gain a firmer indication of sensor performance. These methods, while developed for application in maritime surveillance, also have potential for utility in other areas such as minesweeping and estimation of populations.

Résumé

Le présent rapport expose les résultats du soutien à la recherche opérationnelle dans le cadre d'opérations de surveillance maritime menées dans les approches maritimes du Canada. Le principal résultat est le développement d'une méthode bayésienne servant à estimer le rendement de capteurs, de façon à pouvoir utiliser cette méthode pendant le cours des opérations. Avec cette méthode originale, il n'est pas nécessaire d'avoir une connaissance de la réalité maritime pour évaluer les capteurs. Le rapport présente également une méthode d'évaluation de la vérité mer, qui fournit une capacité d'analyse novatrice et unique pour estimer les brèches de la surveillance. Une série de mesures et de méthodes inédites sont mises en œuvre dans divers outils opérationnels pour analyser la situation maritime (SM), qui appuie des procédures et des processus opérationnels nouveaux. Ces travaux visent à renforcer la capacité de surveillance maritime et à en tirer le maximum.

Importance pour la défense et la sécurité

Le présent rapport expose un ensemble de recherches menées dans le cadre du projet de recherche avancée (PRA) 11hn de Recherche et développement pour la défense Canada (RDDC). Le projet vise à améliorer la surveillance maritime des navires de surface. Les méthodes élaborées se fondent sur l'avantage de disposer de multiples capteurs pour surveiller une région donnée, à un certain moment, afin d'estimer les probabilités de détection à l'aide de capteurs ainsi que le nombre de cibles dans la région qui pourraient être passées inaperçues.

Un avantage complémentaire important de la méthode bayésienne, lorsqu'elle est appliquée à l'analyse de l'ensemble des capteurs, c'est qu'elle peut aussi servir à estimer la réalité maritime (généralement inconnue). Une solution analytique a été produite pour la distribution du nombre total de cibles qui auraient pu se trouver dans les parages dans n'importe quel essai donné d'un système multicapteurs (l'ensemble de toutes les cibles détectées et non détectées). Cette solution suit la loi binomiale négative. Elle permet d'obtenir une estimation utile du rendement de n'importe quel groupe de capteurs employés pour alimenter la situation maritime (SM).

Incidence

Cette recherche a incité les Forces canadiennes à adopter une politique qui préconise l'emploi simultané de multiples capteurs dans les exercices et les essais de surveillance qui ont lieu sur la côte Ouest, afin d'obtenir une indication plus nette du rendement des capteurs. Ces méthodes, bien qu'elles soient développées pour leur application dans la surveillance maritime, pourraient également être utiles dans d'autres domaines, dont le dragage de mines ou l'estimation de populations.

Table of contents

List of figures

Acknowledgements

This report presents a large body of work that is made possible by many people over the years: Andrew Wind, as the scientific lead of ARP 11hn – Maritime Security Planning Tools and Analysis; Pete Smith, who was instrumental in setting up the ARP; and Neil Carson, who helped plant the seeds for many of the tools and concepts presented here. Ron Funk provided many useful comments and corrections. Lt(N) Phil Hagell, as the senior JTFP Surveillance Officer, has provided vigorous support with the recognized maritime picture validation trials.

1 Introduction

1.1 Background

Coastal maritime surveillance is an ongoing Canadian Forces (CF) activity. However, the manner in which maritime surveillance is conducted is ever changing. One of the major changes observed today is the steady increase in the variety of remote sensing capabilities available. For example, RADARSAT 2 (RS2) can detect surface ships on the order of 25 meters in length in a wide area surveillance mode [1]. This is causing a gradual shift away from purely wide-area surveillance using aircraft patrol platforms and more towards remote sensing [2]. This shift in emphasis can be seen in CF publications such as the Maritime Capability Planning Guidance (MCPG) 2010 [3]. To make effective use of these new remote sensing capabilities, it is necessary to gain an understanding of the capabilities and performance of these sensors. The Bayesian approach described in this report was developed in order to support this sensor characterisation.

In order to use new sensor capabilities effectively, it is essential to understand the performance characteristics of that sensor in the operating environment in which it is employed. One of the most important performance metrics is the likelihood that the sensor will detect ships in the area being covered. In other words, if the sensor detects a certain number of ships, how confident can the decision maker be that these are all the ships in the area? A related problem is the possibility that some of the reported targets could be false positives. These issues are especially important if remote sensors are to be used to shoulder an increasing share of surveillance activity, so that the information they provide can be used to form an accurate picture.

The approach that has been undertaken by Operational Research Teams to characterise the sensors is to use operational systems and real-life data. This has the advantage that the data obtained will be representative of future operational use, and data can be collected at relatively low cost. A disadvantage is that analysis of the data may be difficult due to the lack of knowledge of ground truth, and messiness of real-world data. A possible alternative approach would be to set up dedicated trials, but this has its own problems. Apart from the prohibitive cost, a major problem with dedicated trials is that no matter how carefully set-up, they will never perfectly mimic real maritime targets. The amount of data that can be collected from dedicated trials will be limited, both in terms of the types of ships observed, and to limited times and weather conditions.

The difficulty of characterising sensor capability using operational data arises from the fact that our knowledge of sea truth is unknown. This is because our knowledge of sea truth is based on information we receive from sensors, and the specific performance characteristics of these sensors are not well known. This report outlines a method for analysis of operational multi-sensor surveillance data.

As new technology becomes available for surveillance, the processes and procedures to use the technologies also need to be developed or updated to accommodate existing and future systems. This paper discusses recent work by the Defence Research and Development Canada Centre for Operational Research and Analysis (DRDC CORA) in the development of new analysis methods in the field of maritime surveillance. The work described here was done under the auspices of the

DRDC Advanced Research Project (ARP) 11hn – Maritime Security Planning Tools and Analysis.

1.2 Maritime surveillance concepts and terms

Aircraft patrols and remote sensing satellites are only two of the diverse types of sensors/platforms that can contribute to the development of the Recognized Maritime Picture (RMP) on the surface of Canada's ocean approaches. An expanded list would include the following:

a. *Aircraft patrols*, which will attempt to detect (using airborne radar), overfly, and identify each ship in an area of interest (AOI);

b. *Satellite surveillance systems* employing synthetic aperture radar (SAR) sensors, such as RS2;

c. *Shore-based conventional radars*, generally used for traffic control;

d. *Automatic Identification Systems (AIS)*, an automated electronic ship tracking system where an onboard radio transponder broadcasts information on the ship's identification, position, course, etc. which is received by a nearby fixed installation, ship, aircraft, or satellite. Ships of a certain size or type are required by the International Maritime Organization (IMO) to carry AIS transponders;

e. *Long Range Identification and Tracking (LRIT)*, another system mandated by the IMO but somewhat different from AIS, where the ships of a certain size or type are obliged to actively and willingly report via satellite communications their positions at least four times daily. AIS and LRIT are complementary;

f. *Radio signal intelligence gathering systems*, which can potentially extract valuable information from (cooperatively) transmitting targets; and

g. *Surface wave radar (SWR)*, which takes advantage of the transmission properties of high frequency (HF) radio waves hugging the surface of the earth. There is no SWR on the west coast but an experimental system exists on the east coast.

It is important to be able to maximize our understanding of the capabilities of these diverse sensors, and appreciate their relative ability to contribute to the establishment and maintenance of the RMP.

It is useful at this point to pause and define clearly some of the concepts and terms that will be used in this report. Most are broadly used in the field, but some are specific to this report.

a. *The Recognized Maritime Picture* (RMP) is the compilation of available maritime traffic information. The RMP has been described as knowing who is doing what, where and when [5].

b. *Maritime Domain Awareness* (MDA) is the cognitive understanding of a maritime operational area. The RMP is a subset of MDA, as MDA brings in the fifth ‘W’ – why. It includes knowledge of the environment, adversaries, allies, and situational prediction and projection.

c. *A target* is any object of interest (e.g., ships or vessels), which is able to be represented in the RMP.

d. *Knowns* are those targets in the RMP for which the position, classification, and identity are all known.

e. *Detected targets* are those targets in the RMP for which their position is known, but their identity and/or classification are not. There are many sensors (such as radar) that are able to provide target positions, but not necessarily identities.

f. *Undetected targets* are those targets which are not in the RMP. For these targets, there is no positional, classification, or identity information available. This can be due to lack of sensors, ineffective surveillance, or stealth abilities of the target.

g. *Detection* is the activity during which targets are converted from Undetected targets into detected targets. This is the lowest level of information collection for the RMP.

h. *Knowns* are those vessels in the RMP for which the position, classification, and identity are all known.

i. *Identification* is the activity during which targets are converted from undetected targets or detected targets into knowns. Identification is often a prerequisite for a threat analysis.

j. The *Area of Interest* (AOI) is the term used to describe the region for which the CF must maintain an elevated level of surveillance in support of operations.

k. A *Surveillance Session* is defined for the purposes of this analysis as a time and space window during which an AOI is surveyed by all the sensors that can be brought to bear.

l. *Sea Truth* is the underlying reality of what surface ships are actually out there in a given surveillance session. The RMP will always be a subset of the sea truth.

m. *RMP Validation Mission* describes a Surveillance Session established with several key sensors (usually one being an air patrol flight) to update the RMP in an AOI.

n. *Cadence* describes the time periodicity of a sensor - how long it takes to survey the AOI and possibly detect targets that might be out there.

o. Sensor *dwell time* is the time window during which the sensor has the opportunity to detect targets.

p. *Coverage*, that is, how completely is a sensor able to survey an area of interest (AOI) in a given time period.

q. *Time to Detect*, which is the time that the sensor takes to complete a target detection cycle. This could include the time to acquire the sensing data, and process the observations into detected targets.

r. *Probability of Detection* depends on the inherent abilities of the technology employed by the sensor and the physical or behavioural characteristics ('signature') of the target. But it also is dependent on the integrated time duration of this opportunity (a function of time to detect and cadence) and the environmental conditions. In simple terms, it is the probability of the sensor detecting a target when presented with an opportunity to do so.

s. The probability of detection of a target given a time window is often called *detectability*. A target is considered *detectable* if the detectability becomes one for a sufficiently large dwell time.

t. For sensors surveying an area, the sensors will have a *Target Detection Rate (TDR)* which describes the rate at which detectable targets are found.

The methods developed and presented in this report focus on the *probability of detection* by individual sensors of surface ship targets. Useful mathematical formulations to estimate probability values are derived. The focus is on the elemental phenomenon of 'detection', and does not attempt to venture into the complexities of tracking targets over time. Although, the detection and tracking of targets can be related, it is modeled here that the detection is achieved through any means including detect-before-track, or track-before-detect. This report focuses then on the simple question that if there are some targets to be detected in a given AOI within a given time window, then what is the likelihood that the various sensors engaged will detect them?

The mathematics are intended to use actual operational surveillance data accumulated over time to estimate the relative detection capabilities of the various sensors that might be engaged. Bayesian methods are employed, where original 'a priori' engineering estimates of detection performance can be progressively refined and improved over time with the addition of new operational data.

1.3 Scope and aim

In order to present a practical analysis, this report focuses on a systems-level analysis of sensor performance, meaning that the sensors systems are considered as individual components, which feed into a higher level system. The specific physics signal processing, tracking, scan rates, etc. of each sensor are not considered in fine detail. Instead, the output of individual sensors is analyzed. This is not to say that these specifics are not important, but that meaningful results can be obtained by treating sensors as self-contained entities. For example: with RS2, the orbit, look-angle, and detection algorithm are not considered individually. Instead, the target detections are treated as a single output of the sensor.

The overall aim of this report is to present new mathematical methods and procedures applicable to the analysis of maritime surveillance effectiveness.

This report provides the details of these mathematical developments and is intended to complement the broader (and classified) treatment of Maritime Domain Awareness (MDA) presented by the primary author in [4].

1.4 The value of coordinating operational research with operations

In order to use new sensor capabilities effectively, it is essential to understand the performance characteristics of a sensor in the operating environment in which it is employed. Refining estimates of detection probability beyond the level of those provided by the theoretical calculations of the design engineers requires using real-world, empirical, performance data.

Dedicated trials are one way to obtain useful performance data. There are several issues associated with trials. Firstly, there can be very expensive to set up and execute, if not prohibitively so. Secondly, no matter how carefully a trial is set up, it will never perfectly mimic real maritime targets. The amount of data that can be collected from dedicated trials will be limited, both in terms of the types of ships observed, and the variety of times and weather conditions experienced.

Hence, the Operational Research Teams have taken the approach of trying to work with the military operations staff to structure actual operations in a way that might enable 'trial quality' data to be collected. For example, surveillance sessions could be set up where surveillance flights coincide with the presence of other key sensors, such as Radarsat 2 passes, so that the results would provide useful and comparable detection information. This approach has the advantage that the data obtained will be representative of future operational use, and can be collected at relatively low cost. A disadvantage is that analysis of the data may be difficult due to the general 'messiness' of real-world data in comparison to trial data.

Another difficulty associated with characterising sensor capability using operational data arises from not knowing the underlying sea truth. Knowledge of sea truth is limited to the totality of information received from sensors and other sources, and that cannot be expected to be complete. In a Rumsfeldian sense, you can't know what you don't know, so sea truth is impractical (or impossible) to obtain. Hence, the Bayesian mathematics in this report is purposefully developed to extract detection probability estimates when full sea truth is not known. In fact, after performance estimates for the various sensors have been generated, it is even possible to estimate the number of ship targets that were *not* detected in any given surveillance session.

The CF conducts ongoing surveillance activity off all coasts in order to detect potential threats to Canada [23-25]. It is important to be able to measure and assess the performance and effectiveness of these activities. The mathematical methods presented in this report permit surveillance activity to be combined with sensor effectiveness values using software called the Surveillance Analysis Workbook [26].

1.5 Structuring a surveillance session

The surveillance sessions conducted to collect data are conducted in such a manner as to collect data as close to "trial quality" as practicable. These surveillance activities are currently included as part of coastal operational surveillance activities in missions called "RMP Validation Missions". These missions are conducted as follows.

Firstly, an AOI is established within which the surveillance will be conducted. The size of this AOI should be large enough to include sufficient targets for a useful data sample. Conversely, the area should be small enough that all of the sensors included in the trial can survey the entire AOI over the course of the trial. The location of the trial area also needs to be considered. An important factor is the density of targets that can be expected in a given area. An area of very low target density will yield few detection samples for the analysis. On the other hand, in an area of densely congested targets, it will be more difficult to resolve sensor results and to ensure that target detections by different sensors are properly associated to the same target.

The second major consideration is the time window of the surveillance session. It is typically scheduled at a time when the maximum number of sensor platforms will be available to conduct surveillance on the area simultaneously. Figure 1 shows a conceptual depiction of how a surveillance time window is set to capture overlapping sensors. This will typically involve scheduling asset patrols to occur at the same time, for example coordinating an aircraft patrol with the time of a satellite pass. The duration of the session is also critical. It must be set such that it can match the cadence of the slowest sensor. For example, if an aircraft patrol takes one hour to patrol the area, the time window must at least match that hour, even though other sensors may be able to perform many detections during that time. It is also important to keep the duration of the time window fairly constant for several reasons. One reason for this is in order to achieve consistency in the measurement of performance of fast cadence sensors. Another is to match the trial time window to be comparable to typical operational timeframes, which ensures that the sensor performance measured will be relevant for operational use.

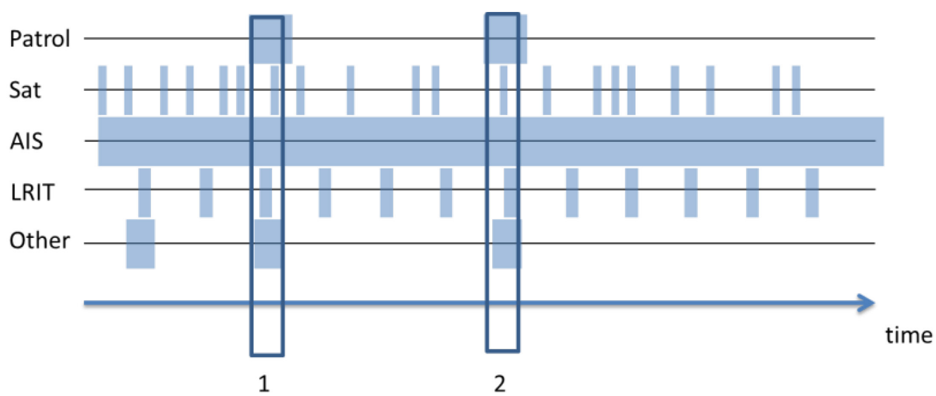

Figure 1: *Example of two possible data collection time windows for multiple sensors in a region of analysis. Shaded areas indicate periods of sensor coverage.*

When the trial is conducted, relevant environmental conditions that could affect sensor performance, such as visibility levels, sea state, and precipitation, should be noted. Then during

the trial, all sensor detection are collected, along with any associated information based on the sensor. This could include information such as time of detection, location, length of target, and any identifying data.

Once all the sensor results are in, the detection from the various sensors are associated manually. The trial operators will then identify all targets in the union of those detected, with the understanding that some targets may have been missed by all sensors. They will have to establish and record which sensors detected which targets. This process can be relatively straightforward if the target density is low and the time window relatively narrow, but may become challenging in high traffic conditions with a longer time window. If data association uncertainties are high enough, based on the subjective call by expert analysts, then the entire session's data should be rejected for use in estimating sensor detection performance. These target detections associated by which sensor detected them will then form the basis for the analysis of sensor detection performance described in this paper.

1.6 Assumptions

Models are, by definition, simplified representations of reality. As such, they entail assumptions. Purely mathematical representations, such as those developed in this report, are abstractions created to enable sufficient simplification that useful analysis can be generated - in this case, useful in providing comparative assessments of the detection abilities of diverse sensors that contribute to the RMP A few fundamental assumptions are clarified first.

The physical operation of the sensor is not explicitly modelled.

Rather, we adopt an empirical approximation of sensor effectiveness, deriving effectiveness coefficients through measurement and observation. A more robust approach could involve physical modelling of the sensors and experimental validation of the models, but that level of effort is outside the scope of this approach.

Sensors are consistent over time.

That is, although sensor performance may depend on various external factors, it will generally perform at the same level tomorrow as it does today. This assumption does not however preclude contributions of variance from changing environmental factors such as noise, weather, etc.

All ships are detectable.

Even though some sensors may not be able to detect a particular ship, there exists no ship which cannot be detected by any sensor. This may seem like a trivial statement, but the mathematics will require it, as ships that are not at all detectable cannot influence the results. Undetectable ships could be considered as a separate category, or "bin", which observations based on sensor detections cannot give us any information about.

Other assumptions are required concerning the practicalities of setting up and using the results from real world surveillance sessions.

Estimated detection probabilities are cumulative values over the entire time window of a surveillance session.

Time is an essential consideration in the calculation of detection probabilities. In general, the longer a sensor looks the more likely it will detect the target. As discussed in the previous subsection, the time window for a surveillance session is established to meet the requirements of the sensor with the slowest cadence. That means sensors with a faster cadence may get multiple opportunities to detect the target ship. As discussed, some effort is made to achieve consistency in these time windows from surveillance session to session. For mobile fast cadence sensors, such as AIS, detection is deemed to occur if a single opportunity is successful during the entire pass of the triggering platform over the target. For fixed fast cadence sensors, such as shore-based radar, detection occurs if just one sweep of the radar detects the target over the entire time the target was within the maximum range of the sensor.

As an aside, the techniques presented in this report do not consider reducing the cumulative probabilities to single detection opportunity values. There are some challenges that must be acknowledged if that was to be done, with an assumption of independence being the primary one. If the opportunities are assumed to be statistically independent then the reduction calculation is straight forward, but the assumption of independence isn't always a good one at the look-to-look level. Conditions influencing detection may be very consistent from opportunity to opportunity, causing the likelihood of detecting in one opportunity very much dependent on whether or not the target was detected in the previous opportunity.

Ship length, ship classification, and sea state are the major considerations in 'binning' surface targets.

It is important to be able to bin ships into categories of (perceived) equal detectability as much as possible. Having too many categories restricts the amount of data in each bin and, hence, the utility of the method. Having too few categories will lead to improper comparisons and incorrect results. This assumption will be discussed in more detail in Section 2.2.

A final assumption must be made about the results of the manual process of associating the inputs from the various sensors engaged in a surveillance session in updating the RMP. There will always be some uncertainties involved. Are two detections by separate sensors looking at the same underlying target in the sea truth, or are they in fact two distinct ships?

All surveillance session uncertainties can be resolved in updating the RMP.

The more congested the traffic is, or the wider the time separation between detections by the different sensors, the more challenging it will be to disentangle the RMP and resolve these differences, and the more untenable this assumption may become. Nonetheless, we have to proceed on the basis that the experienced operators will produce clear and unequivocal detection performance data to feed our calculations. As noted in the previous section, if this assumption is deemed to be untenable for any given surveillance session, then the results should be rejected for use in estimating sensor detection performance.

The purpose of the above set of assumptions is to create the same detection 'playing field' for all sensors that might be engaged, so that they can be fairly compared.

2 Empirical evaluation of maritime surveillance capabilities

One of the fundamentals of sensor performance is its ability to detect a target with given attributes (e.g., physical size, materials), in given environmental conditions (e.g., sea state, cloud cover, day/night), and within a given space and time window. In order to generate a model to describe the sensor performance, the following important point is to be noted:

> The ability to detect a specific target is labelled *probability of detection*, and is defined previously in 1.2.r as the probability of the sensor detecting an unknown target under the given environmental conditions when presented with sufficient opportunity to do so. For the purposes of our model, the definition is simplified to **the probability that there is one or more detection** of a target by a sensor during a surveillance session. This is a valid definition if the integrated sensor TDR divided by the total number of targets is approximately equal to one.

The fundamental challenge of characterising sensor capability using operational data is that analysts will have limited knowledge of sea truth. This is because our knowledge of sea truth is based on information we receive from sensors, and the specific performance characteristics of these sensors are unknown. The mathematical approach developed here aims to overcome this challenge.

2.1 Basic principles employed

The inherent probability of detection associated with sensor A under the given conditions, denoted $P(A)$, will be unknown initially. The system engineers who designed and built the sensor will understand its basic technological capabilities and will have a sense of the anticipated range for the detection probability value. However, $P(A)$ is best estimated empirically – taking the sensor into the field and observing its performance. In a trial where there are N potentially detectable targets and sensor A detects S_A ($\leq N$) of those targets, it is customary and reasonable to use the ratio of these two as an estimate of $P(A)$, denoted $\widetilde{P}(A)$, as shown in equation (1). Of course, it is assumed that in this trial, the sensor dwell time was sufficient such that all detectable targets were registered. Given no additional information, equation (1) provides an unbiased estimate of $P(A)$ based on the trial.

$$\widetilde{P}(A) = S_A / N \tag{1}$$

The application of Bayesian methods permits knowledge about the value of $P(A)$ to be expanded considerably. Empirical observations, combined with any other prior knowledge or evidence of what $P(A)$ might be against similar targets in similar conditions, can be rolled together to produce not only a point estimate of $P(A)$, such as that presented in (1) above, but an estimate of the full probability distribution for $P(A)$. Bayes Theorem [6] will be introduced and employed to develop such distributions.

Previous attempts to evaluate sensor effectiveness rely heavily on using comparisons of the sensor's performance against the 'sea truth', *N*. Practical limitations, however, result in sea truth seldom being available. Note that previously proposed Bayesian methods, such as those in [7], rely on the availability of a sea-truthing sensor.

The methods that are presented here provide a unique view of the evaluation problem. The key to the method proposed is to make observations of the same area of ocean using multiple sensor types within an overlapping time window[1]. By reducing the temporal and spatial uncertainty, the uncertainty in target associations between sensors is greatly reduced. In this situation, the targets detected from the different sensors can then be more easily associated and fused, and then the unique or overlapping contribution from each sensor can be recorded.

It is, of course, vital to note that each sensor is evaluated on the ability to detect targets in time, and no other performance aspects. It is important to also consider the method that the association is done in order to generate the sensor contribution statistics for each target. The following sensor characteristics must be controlled for during the trial:

a. *Sampling rate*: The sensor dwell, and length of the overlapping time window must be sufficiently long such that no sensor is discriminated against for a slow sampling rate. At the same time, a sensor with an extremely high sampling rate may have an unfair advantage. The way that the trial is controlled for this is to plan the sensor collections. In the case when using RS2 in combination with other sensors, the sensor collections would be planned to take place during the pass of the satellite. The window would be long enough to cover the entire area by all sensors and short enough that it would not be covered multiple times.

b. *Sensor coverage*: The area which is surveyed for targets must be overlapping for each sensor. Of course for some sensors, such as those on a patrol asset, it may take longer to cover an area. This is also controlled for, by planning the patrol area and considering the time required to patrol so that the detected targets from the patroller sensors are spatially and temporally overlapping with, say, RS2.

c. *Track hold*: The problem is further simplified by noting that each sensor must detect a target only once. Holding a track on a target is not required for this method as the ability for a sensor to maintain the position of the target is a separate metric from a sensor's ability to detect the existence of a target.

d. *Track identification*: The ability of a sensor to identify a detected target is not required. Since the trial is conducted in a tight spatial and temporal window, the association problem is simplified. Note, however, that there is still a possible issue where two ships in extremely close vicinity to one another are detected by mutually exclusive sets of sensors. In this case, a difficult to detect error of reporting a single target instead of two is possible. However, it is assumed for many applications that this is unlikely to occur. Nevertheless, the possibility when designing the trial should be considered, understood, and controlled for.

---

[1] The length of this time window is chosen such that the dwell time for all of the sensors is sufficiently large such that all targets in the joint field of view are detectable.

From this data, with some analysis of the overlaps between sensors, useful information can be extracted, including individual sensor performance and sensor mix effectiveness. It is even possible to estimate what was not detected.

1. Sensors are assumed to be consistent. That is, although their performance may depend on various external factors, the detection performance of the sensors is consistent over time. The sensor will perform at the same level tomorrow as it does today. This assumption does not however preclude contributions of variance from factors such as noise, weather, etc.

2. Without neglecting the influence of target lengths and target shapes, it is assumed that targets are generally of the same shape, and that targets of similar lengths will have similar effects on sensor detections. This means that targets can be categorized so that one can assume that the contribution from target length within the category is small. For example, the probability of detection with a specific sensor is the same for all “large” ships.

3. All targets are detectable by at least one sensor. Even though some sensors may not be able to detect a particular target, there exists no target which cannot be detected.

2.2 Identifying key variables

There are many different approaches to identifying the key variables that influence $P(A)$. One of the more rigorous of these methods is Factor Analysis [8]. This method uses the statistical correlation between all observable variables to identify any key relationships. This method is used when the relationships between the data are not well understood. Unfortunately, this requires a significant amount of data and would require data collection over an extended period of time. There was not enough data available early on for this type of analysis. The lack of large quantities of data is also a problem for similar methods such as Independent Component Analysis or Cluster Analysis.

As an alternative approach for this analysis, a semi-qualitative method was used to identify the key variables. A quick survey of the literature [9][10][11] reveals that there are some variables that frequently appear in sensor models. The following variables are considered to be the key influences affecting sensor detection performance:

1. Target length or size,

2. Target category (i.e., Merchant, Fishing, or Naval for ships), and

3. Environmental conditions; particularly sea state.

Modelling the sensor detection performance as a function of length ensures that length contributions are captured in the analysis.

Target category affects some sensors more than others. For example, an Automated Identification System (AIS) is internationally mandated for vessels over 300 tons or for passenger vessels. Many fishing vessels are small enough to be exempt (or perhaps are willing to ignore the rules because they do not want to be tracked while fishing).

Influences from environmental conditions are not explicitly modelled in this analysis, but are instead considered as an influence on detection effectiveness variance. However, during the data collection, the environmental influence is controlled by collecting data from all sensors simultaneously over the same area and during the same time.

Following from this discussion on the key variables, the following data should be collected:

- Ship length
- Ship ID
- Ship category
- Set of sensors
- Coverage of sensors
- Environmental conditions

From this set of data, the majority of the effects are expected to be captured which will then be used to calculate the probability of detection.

2.3 A 'naïve' modelling approach

A simple way to sidestep the problem of not knowing sea truth, *N*, is to simply ignore those ships that go undetected. Make the assumption that the union of all sensor observations, *N'*, represents the sea-truth, (i.e., there are no unknowns after observing with all available sensors).

The estimated probability of detection for this approach is simply the ratio of the number of detections by sensor A to the estimated sea truth, *N'*, as presented in equation (2).

$$\widetilde{P}(A) = S_A / N' \tag{2}$$

Note that (2) will tend to overestimate the value of *P*(*A*). The larger the fraction of the population observed (and hence the closer *N'* is to *N*) the better this estimate of the probability of detection will be [12].

The advantage of the naïve method is that it is simple to calculate and does not require a detailed understanding of sensor dependencies. The disadvantage is that it will markedly overestimate detection probability when the combined sensor mix is poor (and hence *N'* is not close to *N*).

Section 3 will describe the alternative statistical modelling approach proposed.

3 Estimating detection probability - A statistical modelling approach

In this approach we assume that sea truth, N, is *not* known. However, we have multiple sensors observing the same area at the same time. We will also make the assumption that the sensors operate independently.

3.1 Estimating detection probability with two sensors

Suppose we have two sensors, A and B, operating simultaneously and independently over the area and time period of interest, and that all targets detected by each sensor are identified. We assume that the detection of a target by each sensor can be modelled as a Bernoulli process so that, given N total targets, the number detected by sensors A and B have independent Binomial distributions with parameters $P(A)$ and $P(B)$ respectively. In executing this trial each sensor detects some number of ships, denoted S_A and S_B, respectively. Some ships are observed as having been detected by both sensors, denoted S_{AB}.

As before, the following estimators can be put forward.

$$\begin{aligned} \widetilde{P}(A) &= S_A / N \\ \widetilde{P}(B) &= S_B / N \\ \widetilde{P}(A \cap B) &= S_{AB} / N \end{aligned} \tag{3}$$

The underlying probability of detection by both sensors can also be usefully expressed as a product involving conditional probabilities and the underlying probabilities of detection by each sensor on its own. Using $P(B|A)$ to represent the conditional probability of detection by sensor B given detection by sensor A, and vice-versa, we know [13] that:

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B \mid A) \\ &= P(B) \cdot P(A \mid B) \end{aligned} \tag{4}$$

The assumption has been made that the sensors are *independent* of one another. The mathematics is always considerably simpler when the independence assumption can be made, so there is a strong motivation to consider that assumption. However, care must be taken to ensure that the independence assumption is justified and reasonable in the context. Otherwise, the mathematics, however elegant it may be, would be incorrect and useless.

The question of independence is one of deciding if one event has an influence on another. Is detection by one sensor dependent on whether another sensor has detected it or not? One would logically expect sensors employing different technologies and platforms to be independent. Hence, the assumption of independence seems reasonable, as long as the sensors under consideration are sufficiently unique.

The expanded phrase *conditional independence* is often applied for clarity (where the 'condition' is that all other factors external to the two sensors in question are equal).

A good example of this is the Automatic Identification System (AIS) and the Long Range Identification and Tracking (LRIT) system. Both sensors report positions of ships using a broadcast. Additionally, both sensors are mandated for large ships (300 gross tonnes and larger) and so report for the same subset of ships. AIS is transmitted using Very High Frequency (VHF) radio signals, whereas LRIT is transmitted using satellite communications. The information reported by the two technologies follows independent channels. If AIS is either not transmitted or not detected, it does not preclude detection from LRIT. The same is true vice versa. In other words, there is no causality between the two sensors and so they are considered to be independent.

It is important to confirm independence between sensors [14]. This can be accomplished computationally from historical data, observing whether two sensors have indeed exhibited independent behaviour. It is well known [12] that independent sensors will exhibit the following property:

$$P(A \cap B) = P(A) \cdot P(B) \tag{5}$$

A complementary report [4] to this report presents a plot showing many historical pair-wise sensor comparisons. The fact that this plot shows the relationship of the number of intersecting detections to the product of the individual detections (when normalized) is roughly linear provides a strong level of confidence in the assumption of independence.

Under independence, conditional probabilities become simply the probabilities themselves. That is:

$$\begin{aligned} P(B \mid A) &= P(B) \\ P(A \mid B) &= P(A) \end{aligned} \tag{6}$$

Under the condition of sensor independence, the third part of Equation (3) can now be updated to reflect the model under independence, where now there are two methods of estimating $\widetilde{P}(A \cap B)$:

$$\begin{aligned} \widetilde{P}(A \cap B) &= S_{AB} / N \\ \widetilde{P}(A \cap B) &= \widetilde{P}(A) \cdot \widetilde{P}(B) = S_A \cdot S_B / N^2 \end{aligned} \tag{7}$$

The problem with equations (3) and (7) is that they contain the sea truth, *N*, which is an unknown quantity. By equating the right hand sides of the two equations above one can solve for *N* and eliminate it from further consideration, while providing an estimate, $\tilde{N}$, for *N* at the same time:

$$\tilde{N} = \frac{S_A \cdot S_B}{S_{AB}} \tag{8}$$

Of course, $\tilde{N}$ will not necessarily be an integer number of ships. This approach for estimating N and its underlying distribution is discussed in greater detail in Section 4. Substituting (8) into (3) now provides estimates for $P(A)$ and $P(B)$ that involve *only* the measured quantities:

$$\begin{aligned} \tilde{P}(A) &= \frac{S_{AB}}{S_B} \\ \tilde{P}(B) &= \frac{S_{AB}}{S_A} \end{aligned} \tag{9}$$

All estimators have specific properties. In this case, it is clear that (8) and (9) are only valid if S_A, S_B, and S_{AB} are non-zero. Note that these results can also be derived by applying the method of moments, equating each observed variable to its expected value.

Equations (8) and (9) present a very useful result. Effectively, this approach substitutes a known subset of targets (those detected by B) for the unknown sea truth N. Then it estimates the performance of A based on how many of this reduced 'sea truth' are detected. The reverse applies for sensor B.

It must be kept in mind how variables such as the target length and the environment can influence the value of the sensor detection performance. Equation (9) does not account for these variables; however, the variables can be controlled in the experiment.

It is important that for every experiment the value for the intersections and total counts be collected under consistent conditions. This is achieved by collecting data over a constrained geographic area and during a small time window. In this way, the contribution from environmental variation is de-convolved from the measurements. Also, the target set will be relatively consistent for all sensors engaged in the experiment.

By measuring the environmental conditions during each experiment one could, with enough data, measure the effect of the environment on the detection performance. Subsequent experiments with similar conditions may be compared and data may even be combined. It would be of uncertain benefit to combine measurements taken during a hurricane with measurements taken on a flat sea.

To control for target variables, the data is sorted by the major target characteristics: ship size and category. The sorted data is analyzed independently providing separate coefficients for each target size/category combination adopted. For example, separate probabilities of detection for large, medium, and small commercial ships can be derived.

Bins representing suitable combinations of environmental conditions and target variables will be created: for example, 'large fishing vessels in high seas', 'medium vessels of all classes in calm

seas', etc. Creating more bins enhances the resolution of the collective estimates. However, it also dilutes the amount of data in each individual bin, which increases uncertainty on the probability estimates for targets in that bin.

3.2 Estimating the uncertainty in sensor detection probability using Bayes' theorem

For the purposes of simplifying equations, the following variable is defined:

$$f_{AB} = S_B - S_{AB} \tag{10}$$

This is the number of failures by sensor *A* on the subset of targets detected by sensor *B*. That is, the number of targets detected by sensor *B* but *not* by sensor *A*. Equation (9) can be re-written as:

$$\widetilde{P}(A) = \frac{s_{AB}}{s_{AB} + f_{AB}} \tag{11}$$

The new 'sea truth' is *X*, where ($X = s_{AB} + f_{AB}$). After *X* target observations, which are treated as a Bernoulli process of independent identical trials each with the probability of detection, *P*(*A*), by sensor *A*, the probability of sensor *A* having *s* detections and *f* failures is given by the binomial distribution [13]:

$$P(s, f \mid P(A)) = \binom{X}{s_{AB}} \cdot P(A)^{s_{AB}} \cdot (1 - P(A))^{f_{AB}} \tag{12}$$

This expression gives the probability of any given result of *s* and *f* after observing *X* targets as a function of the probability of detection by sensor *A*. However, since one is trying to measure *P*(*A*), the inverse of this expression is desired to determine what range of values of *P*(*A*) is likely given the measured *s* and *f*. That is, what is the probability density function for *P*(*A*) given observations of *s* and *f*. One approach to arrive at this distribution for *P*(*A*) is to use Bayesian theory to infer its value.

In a Bayesian analysis there are several terms which are commonly used in a technical sense. The terms 'prior' and 'posterior' refer to the distribution of a random variable before (prior) and after (posterior) any evidence is observed. Bayes' Theorem enables the derivation of the posterior distribution of a random variable given its prior distribution and the observed evidence, which in this case is the number of ship detections by one or more sensors.

As in the previous section, we assume conditional independence for all the sensors.

For a hypothesis, *H*, with evidence, *E*, Bayes' Theorem states [6], [22]:

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)} \tag{13}$$

In this case, the hypothesis is that $P(A) = p$, and the evidence is the observed set $\{ s_{AB}, f_{AB} \}$. $P(H)$ is the prior probability density function of the hypothesis H, $P(E)$ is the marginal probability of observing evidence E for all hypotheses, and $P(E|H)$ is the probability of seeing the evidence E if the hypothesis is correct. In this case, $P(E|H)$ is given by Equation (12).

For the prior probability density function, $P(H)$, there is usually no evidence to make a reasonable guess at what $P(A)$ could be. If one considers any value for $P(A)$ to be equally likely, then $P(H)$ = 1 (See Chapter 8 in [15] for non-informative priors). Finally, $P(E)$ is the a priori probability of witnessing the evidence, which is the sum of the product of the probabilities over all the hypothesis ($\sum P(E \mid H_i) P(H_i)$). The value of *P(A)* given our evidence *s* and *f* is then:

$$P(P(A) = p \mid s_{AB}, f_{AB}) = \frac{\binom{X}{s_{AB}} \cdot p^{s_{AB}} \cdot (1-p)^{f_{AB}} \cdot 1}{\int_0^1 \binom{X}{s_{AB}} \cdot p^{s_{AB}} \cdot (1-p)^{f_{AB}} \cdot 1 \, dp} \tag{14}$$

This can be simplified since the combinatorial function is a constant (i.e., not a function of *p*) in both the numerator and denominator, and hence cancels out of the equation. Also, the integral in the denominator is in the form of a recognized special function – the *Beta function* [16]. The final result for the probability density function for $P(A)$ given evidence from observation is then:

$$P(P(A) = p \mid s_{AB}, f_{AB}) = \frac{1}{B(s_{AB}+1, f_{AB}+1)} \cdot p^{s_{AB}} \cdot (1-p)^{f_{AB}} \tag{15}$$

This is recognized as the *Beta distribution*. The 'best' estimate of $P(A)$ from this is found at the maximum of this function, which is the same as the value from equation (9). Using this distribution, the uncertainty in that estimate can be easily quantified. Figure 2 illustrates the shape of this distribution with $s = 8$ and $f = 2$. As expected, the best estimate of $P(A)$ is 8/(8+2) = 0.8.

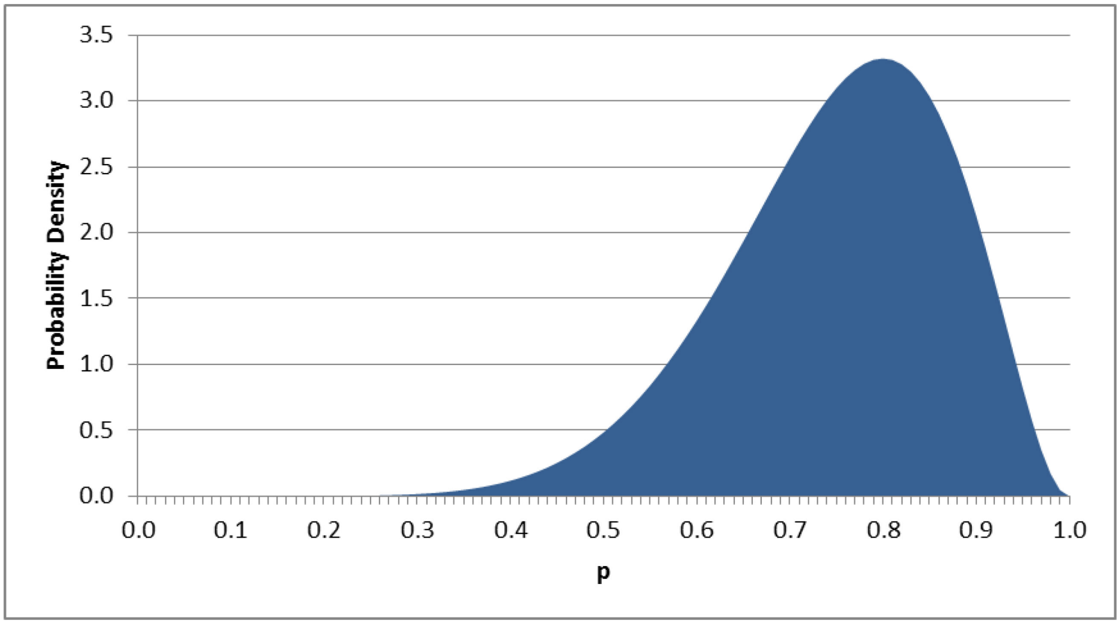

Figure 2: *Plot of the probability distribution for the probability of detection of a sensor given an observation of s=8 and f=2.*

The value of $P(A)$ has been estimated by a first set of measurements, but this process can be iterative. If more data is collected in future experiments, the same Bayesian approach can be used to update the posterior probability of $P(A)$ given additional observations. From each trial or set of observations, the posterior distribution is updated given new observations of successes, s_{AB2}, and failures, f_{AB2}. To do this, Equation (13) is evaluated with $P(H)$ equal to Equation (15):

$$P(\widetilde{P}(A) = p \mid s_{AB}, f_{AB}, s_{AB2}, f_{AB2}) = \frac{\left[\binom{s_{AB2}+f_{AB2}}{s_{AB2}} \cdot p^{s_{AB2}} \cdot (1-p)^{f_{AB2}}\right] \cdot \left[\frac{1}{B(s_{AB}+1, f_{AB}+1)} \cdot p^{s_{AB}} \cdot (1-p)^{f_{AB}}\right]}{\int_0^1 \left[\binom{s_{AB2}+f_{AB2}}{s_{AB2}} \cdot p^{s_{AB2}} \cdot (1-p)^{f_{AB2}}\right] \cdot \left[\frac{1}{B(s_{AB}+1, f_{AB}+1)} \cdot p^{s_{AB}} \cdot (1-p)^{f_{AB}}\right] dp} \tag{16}$$

This evaluates to:

$$P(P(A) = p \mid s_{AB}, f_{AB}, s_{AB2}, f_{AB2}) = \frac{1}{B(s_{AB}+s_{AB2}+1, f_{AB}+f_{AB2}+1)} \cdot p^{s_{AB}+s_{AB2}} \cdot (1-p)^{f_{AB}+f_{AB2}} \tag{17}$$

Equation (17) is an elegant and powerful result meaning that the most current distribution for the probability of detection can be characterized by keeping track of only two variables: successes and failures, which are accumulated over time.

3.3 Generalizing to more sensors

Up to this point only comparisons between two sensors have been discussed. By expanding the definition to multiple sensors, many comparisons can be made. In general, with each of T different sensors one can make T-1 pair-wise comparisons to measure the detection performance. Figure 3 illustrates an example for three sensors.

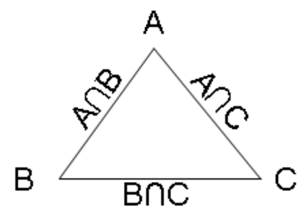

Figure 3: *Diagram illustrating the sensor comparisons for three sensors.*

Using a Bayesian update identical to the way Equation (17) was derived, the probability density for $P(A)$ with comparison to sensors B and C is:

$$P(P(A) = p \mid s_{ab}, f_{ab}, s_{ac}, f_{ac}) = \frac{1}{B(s_{ab} + s_{ac} + 1, f_{ab} + f_{ac} + 1)} \cdot p^{sab+sac} \cdot (1-p)^{fab+fac} \tag{18}$$

Generalizing this result to an arbitrary number of sensors is straightforward:

$$P(P(A) = p) = \frac{1}{B(\Sigma s + 1, \Sigma f + 1)} \cdot p^{\Sigma s} \cdot (1-p)^{\Sigma f} \tag{19}$$

3.4 Advantages and disadvantages of the statistical method

The advantages of the statistical method over the 'naïve' method are:

- It does not require knowledge of sea truth;
- It provides a distribution for unknown detection probabilities. and a measure of uncertainty.

The primary disadvantage of the statistical method is that it requires the coordination of two or more sensors.

4 Estimating sea truth

Using a similar approach, one can estimate the number of missed detections and the distribution of that number. Effectively, one can estimate the sea-truth, N.

Another group has applied a Bayesian method to predict the number of missed mines based on sensor performance [17]. The problem of detecting ships follows the same formulation. In their formulation, they are unable to obtain a closed form solution, and must make approximations due to the method they chose to define their multi-sensor detections[2]. However, it is shown here that by choosing a non-informative Bayesian prior with the multinomial mix, a distribution for estimating the sea truth can be derived in closed form. Additionally, by using the conditional independence assumption, one can also derive a closed-form solution to estimate N without the requirement of a prior estimate of sensor performance.

4.1 The two sensor case

This derivation will be explained starting with the two sensor case, and then generalized for any number of sensors. For two sensors, A and B, which are observing an unknown number of ships, N, there are exactly three measurable sets of observations: ships observed by A (S_A), ships observed by B (S_B), and ships observed by both A and B ($S_A \cap S_B$). From these measurements, for the N ships, each ship can be partitioned into one of four sets:

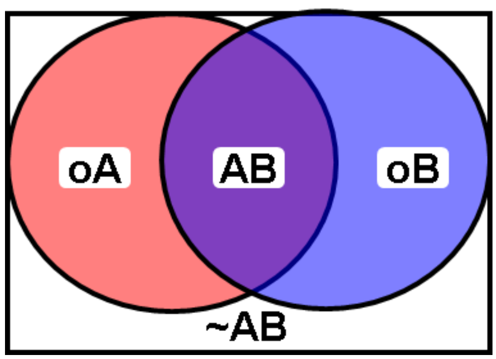

Figure 4: *Partition of sets from observations of two sensors.*

where:

- oA: is only A $= S_A - S_A \cap S_B$
- oB: only B $= S_B - S_A \cap S_B$
- AB: both A & B $= S_A \cap S_B$
- $\sim AB$: neither A nor B $= N - oA - oB - AB$

[2] The paper uses a Dirichlet-Multinomial mixture, which makes the calculation of the marginal probability $P(E)$ difficult.

The sum of the four sets is always going to be equal to the sea truth, which we want to estimate.

$$oA + oB + AB + \sim AB = N, \tag{20}$$

In order to apply the Bayesian inference method, as was done for the measurement of the probability of detection, one needs to describe *P(E|H)* in Bayes formula (13). For the four possible outcomes, the probability that any given subset is observed is described by a multinomial distribution [18]. This distribution is:

$$P(E \mid N, p_a, p_b) = \frac{N!}{oA!oB!AB!\sim AB!} p_{oA}{}^{oA} p_{oB}{}^{oB} p_{AB}{}^{AB} p_{\sim AB}{}^{\sim AB} \tag{21}$$

where the four *p*'s are the probabilities that a given ship falls within the given set, and sum to one. These probabilities can be written in terms of *P*(*A*) (also denoted as p_a) and *P*(*B*) (denoted p_b). Note that:

- $P_{oA} = p_a \cdot (1 - p_b)$,
- $P_{oB} = p_b \cdot (1 - p_a)$,
- $P_{AB} = p_a \cdot p_b$, and
- $p_{\sim AB} = (1 - p_b) \cdot (1 - p_a)$.

At this point, it is important to note that the estimates for p_a and p_b cannot come solely from the current experiment for which we are trying to estimate *N*. That is, we cannot substitute (9) for p_a and p_b. The reader will note the circularity in estimating p_a and p_b from observed detections (with other non-detections unknown but implicit) and then turning those values around to estimate non-detections. The estimates for p_a and p_b must come from previously derived information.

Continuing, equation (21) can then be re-written (see details in Annex A) as:

$$P(E \mid N, p_a, p_b) = \frac{N!}{oA!oB!AB!\sim AB!} p_a{}^{S_A} p_b{}^{S_B} (1 - p_a)^{N - S_A} (1 - p_b)^{N - S_B} \tag{22}$$

Returning to Bayes' Theorem and equation (13), *P*(*E*|*H*) is well defined by (22).

The fraction *P*(*H*)/*P*(*E*) is the next part that needs to be defined. *P*(*H*) is the a priori probability that the hypothesis is correct, and *P*(*E*) is the sum $\sum P(E \mid H_i)P(H_i)$ over all possible hypotheses.

$$\frac{P(H)}{P(E)} = \frac{P(N=n)}{\sum_{i=0}^{\infty} P(E \mid N=i) \cdot P(N=i)} \tag{23}$$

This type of summation is generally difficult for estimating a hypothesis where N can be any number from 0 to infinity. If one has some information that would indicate the likely range of values of N, an informative prior distribution could be defined. In the absence of such information one can use a non-informative prior. Although there is some philosophical debate on the use and form of non-informative priors [19], it is argued here that without prior information, assuming all possible values of N are equally probable - a uniform distribution - is reasonable. Another accepted non-informative prior is a function that decreases in proportion to $1/N$, which despite seeming 'informative' relies on the argument that very large values of N must be less likely. With either choice, the risk of choosing a mis-specified prior (very different than the distribution one is measuring), which can have a major effect on the outcome [20], should be avoided.

Also, it should be noted that both the uniform or reducing (proportional to $1/N$) priors are improper distributions, as they cannot be made to sum to 1 unless an upper bound on N is assumed. Because of the mathematics of the Bayesian analysis an improper prior causes no difficulties in the final calculations.

It will be shown that by choosing the uniform distribution as a prior a closed form solution for $P(E)$ can be determined.

Since during an experiment one observes $S_A + S_B - S_{A \cap B}$ ships in total, it is not possible that N is less than this value[3]. This means that the summation does not have to start at zero. Using this information, and that the probability of any N above the number of ships observed is equally probable, the summation becomes:

$$\frac{P(H)}{P(E)} = \frac{1}{\sum_{i=S_A+S_B-S_{A\cap B}}^{\infty} P(E \mid N=i)} \tag{24}$$

Substituting in Equation (22) and performing the summation (see Annex A), one finds that the value converges to:

$$\sum_{i=S_A+S_B-S_{A\cap B}}^{\infty} P(E \mid N=i) = \frac{(oA+oB+AB)!}{oA!\,oB!\,AB!} \cdot \frac{p_a \cdot p_b \cdot (1-p_a) \cdot (1-p_b)}{\left[1-(1-p_a)\cdot(1-p_b)\right]^{1+oA+oB+AB}} \tag{25}$$

[3] This is assuming no false detections. The effect of false detection will be discussed at the end of the section.

Putting Equation (22) and (25) together (see Annex A) and simplifying results in an elegant solution:

$$P(N=n\,|\,E)=\binom{n}{S_A\bigcup S_B}P(\sim D)^{N-S_A\bigcup S_B}\cdot\left[1-P(\sim D)\right]^{S_A\bigcup S_B+1} \tag{26}$$

Where $P(\sim D)$ is the probability of not detecting a target ship defined as $P(\sim D) = (1\text{-}p_a)(1\text{-}p_b)$. This is recognizable as the Negative Binomial distribution [13].

The Negative Binomial distribution naturally represents the case of estimating the number of independent Bernoulli trials (with given probability, p, of success) required to produce a fixed number, s, successes, denoted $NB(s,p)$. While similar, our detection problem has a distinct difference: one does not know when a failure has occurred (i.e., a ship goes undetected). The solution (26) is distributed as $NB(\ S_A \bigcup S_B + 1,\ P(\sim D)\ )$, which would be the distribution of the number of trials required (i.e., the sea truth) to produce $S_A \bigcup S_B + 1$ successes if one knows when a failure has occurred. The fact that one does not know when failures occur seems to introduce the requirement for an extra 'pseudo-success'. Future work (and the reader might like to help out) will try to define the probabilistic logic that supports this "plus 1" phenomenon.

Interestingly, another group has arrived at a Negative Binomial distribution using different distributions and priors for detecting submarines [19]. Their method discussed the prediction of the number of submarines one would expect to detect given some previous observations. Their formulation discusses a measurement of the expected number of submarines which will be observed on a given patrol using information from multiple previous observations (a method which can be considered analogous to repeated observations by sensors). They defined detection probabilities in terms of a rate of submarine detections in a Poisson process[4] instead of the sensor's probability of detection. However, the end result was still the same distribution. Their approach is effective when the chance of detecting a ship for any given observation or patrol is small. In the case of the modern RMP, ship detections are more frequent, and so the method used here can be applied.

4.2 Extending to three or more sensors

The elegance of (26) is striking. It involves only the probability of *any* sensor detecting the target or not. The total number of sensors involved is not really a factor at all. This permits the same expression to be employed for three or more sensors:

$$P(N=n)=\binom{n}{\bigcup S_i}P(\sim D)^{N-\bigcup S_i}\cdot\left[1-P(\sim D)\right]^{\bigcup S_i+1} \tag{27}$$

[4] That is, the longer you observe, the more likely you will detect a submarine.

Equation (27) is straightforward to compute, and provides an estimate of N including confidence intervals. For example, given three sensors with probabilities of detection of 0.4, 0.5, and 0.8, and one observes $\bigcup S_i = 30$ ships, the estimation for N can be illustrated as shown in Figure 5.

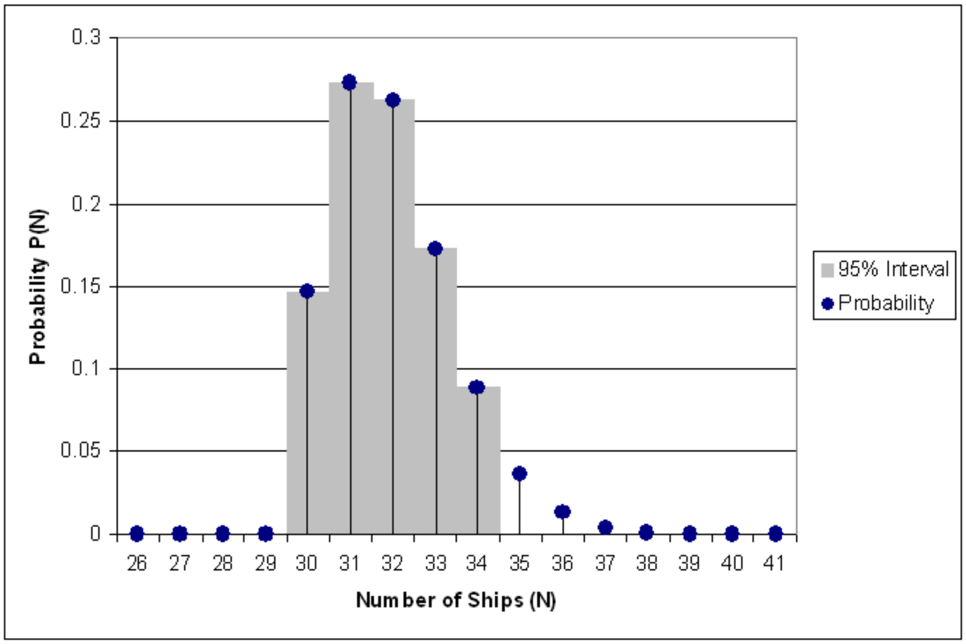

Figure 5: *Estimation of missed ships given 30 detections and 3 sensors p=0.4, 0.5, and 0.8.*

The interpretation of Figure 5 is that there is most likely to be 31 ships (one missed), and with 95% confidence, there were zero to four ships not detected.

Mr. Emond, one of the co-authors, has developed a Monte Carlo simulation of the broader Bayesian analysis. This tool is described in detail in Annex B, along with a detailed example. This numerical approach permits the use of any desired prior distributions for N, any desired number of sensors, and Beta-distributed prior distributions for each sensor's detection probabilities, $P(A)$, $P(B)$, $P(C)$, etc. It then numerically calculates the posterior distribution for N based on the number of individual ship detections observed. Mr. Emond would be pleased to make this software available to any interested readers.

The example in Annex B.3 gives excellent insight into how assumptions on prior distributions will influence the shape of the posterior distribution.

4.3 Impact of false targets

False targets have not been considered in the analysis to this point. If a sensor generates false detections, then these are 'targets' that other sensors (presumably) would not detect. The estimates of sensor performance of other sensors will be reduced somewhat and the estimate for undetected targets must increase.

False alarm rates for sensors are usually well understood and assumed to be low. In any event, they can probably be measured. Repeating the derivation considering false alarm rates of sensors would allow for adjustment of the estimation to account for typical sensing conditions.

The impact of false alarm rates on the distribution given in Equation (27) is, however, well understood. The mean and variance of the distribution are:

$$\mu_n = \left(\bigcup S_i + 1\right) \cdot \frac{P(\sim D)}{1 - P(\sim D)}, \ \sigma_n{}^2 = \left(\bigcup S_i + 1\right) \cdot \frac{P(\sim D)}{\left(1 - P(\sim D)\right)^2} \tag{28}$$

Incrementing $\bigcup S_i$ by one (one sensor generated a false detection) will increase the mean number of missed targets by an amount $P(\sim D) / (1 - P(\sim D))$ and the variance by $P(\sim D) / (1 - P(\sim D))^2$. Note that if the total sensor quality degrades, the impact of false alarms will increase.

To mitigate the impact of false alarms, a simple statistical test to detect false alarms can be defined. Since a false alarm impacts only one sensor, one can test the occurrence of ships detected by only that one sensor and not by any of the others. This permits the flagging of this subset of detections as likely to be including one or more false targets.

5 Conclusion

This report has presented and discussed a body of work under DRDC ARP 11hn aimed at improving maritime surveillance. The key conclusions and main points of discussions are summarized here.

The methodologies developed here for application in maritime surveillance should also find utility in other areas, such as minesweeping and estimation of populations.

5.1 Estimating detection probability using a multi-sensor approach

Any individual sensor has its detection performance limited by the technology it employs, the characteristics of the target, and the environmental conditions at the time. But given a set of conditions, detection of surface shipping on the open seas will always be very much a random phenomenon. The best way to determine a sensor's operational capability is to observe performance empirically in trials. The biggest problem, however, is that one does not know when a detection failure has occurred. Other than in very well controlled trials, sea truth generally will not be known.

This research has demonstrated the value of using multiple sensors in estimating sensor performance. The logic is to take the subset of targets detected by one sensor as a substitute for sea truth and examine how the other sensors perform on that subset. Under the assumption of sensor independence, argued here as not being unreasonable, one can generate simple point estimates of each sensor's performance.

Further, the application of Bayes' Theorem permits the distribution of these estimated probabilities to be calculated as well. A closed form solution has been derived in this report, which is in the form of the Beta distribution.

When employing this Beta distribution as the prior distribution for detection probability in a subsequent trial, the recalculated solution is also a Beta distribution but with parameters adjusted to reflect the new sensor performances. This elegant result permits the accumulation of detection opportunities and successes as trials with similar conditions occur over time to estimate detection probability (and its distribution).

5.2 Estimating sea truth

A major side-benefit of the Bayesian approach when applied to sensor mix analysis is that it can also be used to estimate the (generally unknown) sea-truth. A closed form solution has been derived for the distribution of the total number of targets that may have been out there in any given multi-sensor trial (the union of all detections plus the undetected targets). This solution is in the form of the Negative Binomial distribution. This provides a very useful estimate of the performance of any given collection of sensors employed to feed the RMP.

One of the elegant features of the Negative Binomial solution is that the calculation for the distribution of total number of targets is equally simple for any number of sensors being employed. The key detection probability employed is the probability of *all* sensors not detecting a given target.

This closed form solution relies on the assumption of a uniform distribution as the non-informative prior distribution for the sea truth variable, N. A numerical Bayesian analysis tool is also presented in this report for use with any form of more informative prior distribution for N.

5.3 Operational use

The purpose of developing this theoretical framework for analyzing multi-sensor detection performance is to support operational missions of the Canadian Forces. The Bayesian method has had an operational impact in two primary ways: changing operational processes to permit data collection to support analysis, and then using the results of the analysis to support command surveillance goals.

In order to allow for more effective collection of sensor performance data, the mission profiles of surveillance missions have been modified to facilitate data collection. This has been done by creating a procedure for a type of patrol called an "RMP Validation Mission". This mission involves coordinating an aerial patrol covering a certain area with as many other sensors simultaneously collecting data from the same area. In this way, data can be collected which can then be analyzed using the Bayesian method to ascertain the effectiveness of each sensor.

The CF conducts ongoing surveillance activity off all coasts in order to detect potential threats to Canada [23-25]. It is important to be able to measure and assess the performance and effectiveness of these activities. The mathematical methods presented in this report permit surveillance activity to be combined with sensor effectiveness values using software called the Surveillance Analysis Workbook [26].

This page intentionally left blank.

References

[1] RADARSAT-2 Product Description RN-SP-52-1238, August 2010, MacDonald Dettwiler and Associates Ltd.

[2] Minutes of the 6th Intelligence, Surveillance and Reconnaissance (ISR) Working Group (WG) Meeting, 19 November 2009.

[3] Maritime Capability Planning Guidance (MCPG), 2010.

[4] Horn, S., Unpublished report.

[5] Dickinson, R.G., Martel, S., Jesion, A., Measures of Performance and Effectiveness for Surface Surveillance, ORD Project Report PR1/97, August 1997.

[6] Berger, J.O., Statistical Decision Theory and Bayesian Analysis, Second Edition, Springer-Verlang New York Inc., 1980.

[7] Gauthier, Y., Minter, P., Multi-Source Data Fusion for the Estimation of Mean Shipping Densities, DRDC CORA TM 2005-18, June 2005.

[8] Cureton, E.E., D'Agostino, R.B., Factor Analysis: An Applied Approach, Lawrence Erlbaum Associates, Inc., 1983.

[9] Valin, P., Jouan, A., Gagnon, L., Bossé, É., Airborne Fusion of Imaging and Non-Imaging Sensor Information for Maritime Surveillance, Proc. International Conference on Quality Control by Artificial Vision, Trois-Riviéres, pp. 281-286, May 1999.

[10] Gagnon, L., Klepko, R., Hierarchical Classifier Design for Airborne SAR Images of Ships, Proc. SPIE #3371 conference on Automatic Target Recognition VIII, April 1998.

[11] Valin, P., Bossé, É., Jouan, A., Airborne Application of Information Fusion Algorithms to Classification, DRDC Valcartier Technical Report TR 2004-282, May 2006.

[12] Johnson, R.A., Bhattacharyya, G. K., Statistics: Principals and Methods, 6th edition, John Wiley and Sons, 2010.

[13] Bulmer, M.G., Principles of Statistics, Second Edition, Dover Publications Inc., 1979.

[14] Nasser, R., Deciding Whether There is Statistical Independence or Not?, J. Math. & Stat., 3 (3) pp. 151-172, 2007.

[15] Bolstad, W.M., Introduction to Bayesian Statistics, Second Edition, John Wiley and Sons, 2007.

[16] Rosenkrantz, W.A., Introduction to Probability and Statistics for Science, Engineering, and Finance, Taylor & Francis Group, LLC, 2009.

[17] Bryan, K., Carthel, C., A Bayesian approach to predicting an unknown number of Targets Based on Sensor Performance, Proc. 9th International Conference on Information Fusion, Florence, IT, 2006.

[18] Chiang, C.L., Statistical Methods of Analysis, World Scientific Publishing Co. Pte. Ltd., 2003.

[19] McCue, B, A Justification of a Negative Binomial Model For Target Sightings, MORS, V8 N4, 2003.

[20] Kelley, W.M., The Complete Idiot's Guide to Calculus 2nd Edition, Alpha Books, 2002.

[21] Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., Bayesian Data Analysis, Chapman and Hall, UK, 1995.

[22] Commander Canada Command: Operation Order 10202/09.

[23] Maritime Domain Awareness: Concept of Operation Within MARLANT Area of Responsibility (AOR), June 2007.

[24] Commander Joint Task Force (Pacific) Operation Plan 12100/07.

[25] Wind, A., Horn, S., *Surveillance Analysis Workbook (SAW) Version 11.2,* DRDC CORA TM 2009-039, October 2009.

Annex A Derivation of the closed-form Bayesian multisensor estimate of missed detections

This annex describes in more detail the derivations presented in section 4.

A.1 The conditional probability

For two sensors, *A* and *B*, the probability of observing a given number of ships was defined by the multinomial distribution:

$$P(E \mid N, p_a, p_b) = \frac{N!}{oA!oB!AB!\sim AB!} p_{oA}{}^{oA} p_{oB}{}^{oB} p_{AB}{}^{AB} p_{\sim AB}{}^{\sim AB} \tag{A.1}$$

Where the following identities are defined:

- *oA*: only *A* $= S_A - S_A \cap S_B$
- *oB*: only *B* $= S_B - S_A \cap S_B$
- *AB*: both *A* & *B* $= S_A \cap S_B$
- *~AB*: neither *A* nor *B* $= N - oA - oB - AB$
- Probability of only *A*: $p_{oA} = p_a \cdot (1 - p_b)$
- Probability of only *B*: $p_{oB} = p_b \cdot (1 - p_a)$
- Probability of *A* & *B*: $p_{AB} = p_b \cdot p_a$
- Probability of neither *A* nor *B*: $p_{\sim AB} = (1 - p_b) \cdot (1 - p_a)$

Substituting the probability values into last four multiplicands and their exponents in equation (A.1) gives:

$$\begin{aligned} & p_{oA}{}^{oA} p_{oB}{}^{oB} p_{AB}{}^{AB} p_{\sim AB}{}^{\sim AB} = \\ & \left[p_a(1-p_b)\right]^{S_A - S_A \cap S_B} \cdot \left[p_b(1-p_a)\right]^{S_B - S_A \cap S_B} \cdot \ldots \\ & \left[p_b p_a\right]^{S_A \cap S_B} \cdot \left[(1-p_b)(1-p_a)\right]^{N - S_A - S_B + S_A \cap S_B} \end{aligned} \tag{A.2}$$

Expanding out and collecting exponents gives:

$$p_{oA}{}^{oA} p_{oB}{}^{oB} p_{AB}{}^{AB} p_{\sim AB}{}^{\sim AB} =$$
$$p_a{}^{S_A - S_A \cap S_B + S_A \cap S_B} p_b{}^{S_B - S_A \cap S_B + S_A \cap S_B} \cdot \ldots \tag{A.3}$$
$$(1-p_a)^{S_B - S_A \cap S_B + N - S_A - S_B + S_A \cap S_B} (1-p_b)^{S_A - S_A \cap S_B + N - S_A - S_B + S_A \cap S_B}$$

Simplifying the above results in:

$$p_{oA}{}^{oA} p_{oB}{}^{oB} p_{AB}{}^{AB} p_{\sim AB}{}^{\sim AB} = p_a{}^{S_A} p_b{}^{S_B} \cdot (1-p_a)^{N-S_A} (1-p_b)^{N-S_B} \tag{A.4}$$

Finally, re-substituting into Equation (A.1) gives:

$$P(E \mid N, p_a, p_b) = \frac{N!}{oA!oB!AB!\sim AB!} p_a{}^{S_A} p_b{}^{S_B} (1-p_a)^{N-S_A} (1-p_b)^{N-S_B} \tag{A.5}$$

A.2 The marginal probability

To evaluate the marginal probability, one must sum Equation (A.5) over all possible values of N. Given the set of observations from the sensor, the minimum value of N is taken to be the union of the ships observed by all sensors, $\cup S_i$, which will be represented by the variable U for clarity. The sum will run from U to infinity:

$$P(E) = \sum_{N=U}^{\infty} \frac{N!}{oA!oB!AB!\sim AB!} p_a{}^{S_A} p_b{}^{S_B} (1-p_a)^{N-S_A} (1-p_b)^{N-S_B} \tag{A.6}$$

To evaluate this infinite sum, first, the constants are pulled out of the summation. Also, note that $\sim AB = N - U$. The equation becomes:

$$P(E) = \frac{p_a{}^{S_A} p_b{}^{S_B} (1-p_a)^{-S_A} (1-p_b)^{-S_B}}{oA!oB!AB!} \cdot \sum_{N=U}^{\infty} \frac{N!}{(N-U)!} \left[(1-p_a)(1-p_b)\right]^N \tag{A.7}$$

To simplify this equation, note that the product $[(1-p_a)(1-p_b)]^N$ is, in fact, the probability of not detecting a target, which will be denoted as $P(\sim D)$. In addition, a change of variables is introduced where $k = N - U$. The range of integration is now from $k = 0$ to infinity. To de-clutter the equation further, the terms outside the summation will be denoted by a variable C.

$$P(E) = C \cdot \sum_{k=0}^{\infty} \frac{(k+U)!}{(k)!} P(\sim D)^{k+U} \tag{A.8}$$

Next, the factorials can be re-written as a binomial coefficient by noting that:

$$\frac{(k+U)!}{(k)!} = \frac{(k+U)!}{(k)!}\frac{(U)!}{(U)!} = \binom{k+U}{k} \cdot U! \tag{A.9}$$

Pulling the constants out of the sum, and substituting in Equation (A.9) gives:

$$P(E) = C \cdot U! \cdot P(\sim D)^U \cdot \sum_{k=0}^{\infty} \binom{k+U}{k} P(\sim D)^k \tag{A.10}$$

The sum can be evaluated by noting the Taylor [20] expansion of the following function centered on zero[5].

$$\frac{1}{(1-x)^{U+1}} = 1 + (U+1) \cdot x + \binom{U+2}{2} \cdot x^2 + \cdots = \sum_{k=0}^{\infty} \binom{k+U}{k} \cdot x^k \tag{A.11}$$

Using Equation (A.11), the infinite series from (A.10) is straightforward to evaluate resulting in the following solution:

$$P(E) = \frac{C \cdot U! \cdot P(\sim D)^U}{\left[1 - P(\sim D)\right]^{U+1}} \tag{A.12}$$

Re-substituting the constant results in:

$$P(E) = \frac{U!}{oA!oB!AB!} \cdot \frac{p_a^{S_A} p_b^{S_B} (1-p_a)^{U-S_A} (1-p_b)^{U-S_B}}{\left[1 - P(\sim D)\right]^{U+1}} \tag{A.13}$$

A.3 The posterior probability

$$P(N = n \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)} \tag{A.14}$$

Using Equations (A.13), (A.1) and (13), the posterior probability is:

[5] This is the Maclaurin Series, or just power series.

$$P(N=n\,|\,E)=\frac{\dfrac{n!}{oA!oB!AB!\sim AB!}p_a{}^{S_A}p_b{}^{S_B}(1-p_a)^{n-S_A}(1-p_b)^{n-S_B}}{\dfrac{U!}{oA!oB!AB!}\cdot\dfrac{p_a{}^{S_A}p_b{}^{S_B}(1-p_a)^{U-S_A}(1-p_b)^{U-S_B}}{\left[1-P(\sim D)\right]^{U+1}}} \tag{A.15}$$

This is simplified by noting that $\sim AB! = (N-U)!$ and that $P(\sim D) = (1-p_a)\cdot(1-p_b)$:

$$P(N=n\,|\,E)=\frac{\dfrac{n!}{oA!oB!AB!(N-U)!}p_a{}^{S_A}p_b{}^{S_B}P(\sim D)^n}{\dfrac{U!}{oA!oB!AB!}\cdot\dfrac{p_a{}^{S_A}p_b{}^{S_B}(1-p_a)^{U}(1-p_b)^{U}}{\left[1-P(\sim D)\right]^{U+1}}} \tag{A.16}$$

This is re-arranged and simplified to:

$$P(N=n\,|\,E)=\frac{n!}{U!(n-U)!}\cdot P(\sim D)^{n-U}\cdot\left[1-P(\sim D)\right]^{U+1} \tag{A.17}$$

The resulting closed-form solution, which is the same as Equation (27), is:

$$P(N=n\,|\,E)=\binom{n}{U}\cdot P(\sim D)^{n-U}\cdot\left[1-P(\sim D)\right]^{U+1} \tag{A.18}$$

Annex B A computational Bayesian model for number of targets in a multi-sensor environment

Let N be the total number of targets present in the area/time window. In a Bayesian statistical model, we treat N as a random variable with distribution function $f(N)$ whose domain is the set of non-negative integers: $\{ 0, 1, 2, \ldots \}$. The aim of a Bayesian analysis is to derive the posterior distribution of the random variable N given the evidence of the number of detections by one or more sensors using Bayes' Theorem.

In a Bayesian analysis there are several terms which are commonly used in a technical sense. The terms prior and posterior refer to the distribution of a random variable before (prior) and after (posterior) any evidence is observed. Bayes' Theorem enables the derivation of the posterior distribution of a random variable given its prior distribution and the observed evidence which in this case is the number of target detections by one or more sensors.

Another term used in a Bayesian analysis is hierarchical model. In a hierarchical model one or more unknown parameters required for the main model are treated as random variables and given prior distributions in order to allow calculation of the random variables of interest in the main model. This type of model is also referred to as a complete Bayesian analysis. Reference [22] is a recommended source for information on Bayesian modelling.

B.1 Mathematical preliminaries

As in the previous section, we assume conditional independence for all the sensors. That is, given N total targets in the area/time window, the number detected by each sensor is independent of the number detected by any of the others. We also assume that the targets are sufficiently similar that the detection of targets by any sensor can be treated as a Bernoulli process of independent identical trials, each with a fixed probability of success.

Let the number of independent sensors be denoted by K where K is a positive integer. Let the number of detections by sensor i be denoted by S_i where S_i is a non-negative integer value.

Given the values S_1 , S_2 , … , S_K we wish to estimate the total number of targets N in the area and time of interest. First we note that N must be at least as large as the largest number of detections by any of the K sensors. If the sensors provide target identification, then the constraint on N is given by the number of unique targets detected by the K sensors. In either case, we denote the minimum value of N by N_{min} where N_{min} is a non-negative integer value. Note that just as the pattern of individual detections by the K sensors provides a way to estimate N (as shown in the previous section), it enters into the Bayesian analysis both by determining N_{min} and by way of the calculation of the posterior distribution using Bayes' Theorem as shown below. However the overlapping identification of individual targets does not otherwise influence the resulting posterior distribution in a Bayesian analysis due to the assumptions of conditional independence between sensors and Bernoulli trial detections by each sensor.

B.2 The Bayesian model

Under the assumption of Bernoulli trials for each sensor, we can write the probability density function for the number of targets S_i detected by sensor i as a Binomial distribution as follows. The variable P_i is the probability of detection of an individual target by sensor i.

$$f(S_i \mid N) = Const.* P_i^{S_i} * (1 - P_i)^{N-S_i} \quad for\ S_i = 0,1,...,N \tag{B.1}$$

The likelihood function for the observations S_1 , S_2 , … , S_K may then be written as follows.

$$L(S_1, S_2, \ldots, S_K \mid N) = Const.* \prod_{i=1}^{K} P_i^{S_i} * (1-P_i)^{N-S_i} \tag{B.2}$$

If we denote the prior distribution of N as $\pi(N)$, we can use Bayes' Theorem to write the posterior distribution of N given the evidence S_1 , S_2 , … , S_K as follows. (A detailed discussion of the prior distribution will be given below).

$$g(N \mid \mathrm{S}_1, S_2, \ldots, S_K) = Const.* \pi(N) * \prod_{i=1}^{K} P_i^{S_i} * (1-P_i)^{N-S_i} \quad for\ N = N_{\min}, \ldots \tag{B.3}$$

The prior distribution $\pi(N)$ represents the known distribution of the number of targets present in the area/time window in a general way, before any information from the sensors is considered. In many cases a non-informative prior is used either because no information is available or because a conservative estimate of the distribution of N is required. If prior information about N is available either from other sources or from historical data, this can be incorporated into $\pi(N)$ directly. This is a major advantage of a Bayesian analysis.

The non-informative prior for the random variable N is given below.

$$\pi(N)\ \alpha\ \frac{1}{N} \quad for\ N = N_{\min}, \ldots \tag{B.4}$$

Note that this prior is improper because it cannot be made to sum to 1 unless an upper bound on N is assumed. Because of the mathematics of the Bayesian analysis an improper prior causes no difficulties in the final calculations.

Given a specific form for $\pi(N)$ either from past data or as a non-informative prior, we could use the above equations to calculate the posterior distribution of N given the evidence S_1 , S_2 , … , S_K, assuming that the detection probabilities P_1 , P_2 , … , P_K are known. If these probabilities are not known, we proceed to apply a complete Bayesian analysis by incorporating these values as random variables and integrating over them to find the required posterior distribution of N.

Incorporating the detection probabilities as random variables

We consider the individual sensor detection probabilities to be independent random variable in the Bayesian sense, each with a prior distribution $Z_i(P_i)$. As above, if specific information about P_i is known it can be incorporated directly into $Z_i(P_i)$. Otherwise a non-informative prior can be

used. In this case the non-informative prior is the uniform or rectangular distribution on the interval (0,1).

While a non-informative prior is often preferred for N, it is less desirable for the detection probabilities. One of the most popular informative prior distributions for detection probability is the Beta distribution.

Beta prior distribution for the sensor probability of detection

As noted earlier, we make the assumption that given N targets in the area and time interval of interest, the number of targets detected by sensor i can be modeled as a Binomial distribution with parameter P_i.

$$f(S_i \mid N) = Const. * P_i^{S_i} * (1-P_i)^{N-S_i} \quad for\ S_i = 0,1,\ldots,N \tag{B.5}$$

Suppose that from a previous exercise where the number of targets, M (also known as ground truth) was known, we have the information that m targets were detected. If we assume a uniform (non-informative) prior distribution for P_i then using Bayes' Theorem we can write the posterior distribution of P_i as follows.

$$f(P_i \mid m\ out\ of\ M) = Const. * P_i^{m} * (1-P_i)^{M-m} \quad where \quad 0 \le P_i \le 1 \tag{B.6}$$

The above distribution is a Beta distribution with parameters (α,β) equal to (m+1, M-m+1).

Even in the case where no information from a prior exercise is known, it is convenient to use the Beta distribution as the default prior for P_i due to its ease of use mathematically and because its interpretation in terms of previous experience is clear. It is particularly helpful if both of the parameters of the Beta prior are integers. This fits in well with the interpretation of previous trial experience with the sensor.

The form of the Beta distribution usually used in statistics in terms of its parameters α and β is given below.

$$f(P_i) = Const. * P_i^{\alpha-1} * (1-P_i)^{\beta-1} \quad where \quad 0 \le P_i \le 1 \tag{B.7}$$

The mean, variance and mode of the Beta distribution are given below.

$$\begin{aligned} E(P_i) &= \frac{\alpha}{\alpha+\beta} = \frac{m+1}{M+2} \\ Var(P_i) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \end{aligned} \tag{B.8}$$

$$Mode(P_i) = \frac{\alpha-1}{\alpha+\beta-2} = \frac{m}{M} \tag{B.9}$$

Posterior distribution of the number of targets - complete Bayesian model

We are now in a position to use Bayes' Theorem to write the posterior distribution of the number of targets N in the area and time of interest given $S_1, S_2, \ldots, S_K$ detections by the K sensors. The general form is given below where we write the prior distribution of N as $\pi(N)$ and the prior distributions of the K sensor detection probabilities as $\pi_i(P_i)$ for $i = 1, 2, \ldots, K$.

$$g(N \mid S_1, S_2, \ldots, S_K) \;\alpha\; \int_{P_1}\int_{P_2}\cdots\int_{P_K} \pi_1(P_1)\ldots\pi_K(P_K)\,\pi(N) \prod_{i=1}^{K} \binom{N}{S_i} P_i^{S_i}(1-P_i)^{N-S_i}\, dP_1 \ldots dP_K$$

$$for\ N = N_{\min}, N_{\min}+1, N_{\min}+2, \ldots \tag{B.10}$$

The above formula appears daunting at first glance but in fact it can be readily evaluated numerically for all values of N by using Monte Carlo integration. This technique is self-weighting in that values of P_i are chosen randomly from their respective distributions rather than evaluating all possible combinations. Note that having evaluated the above formula for all values of N up to some upper limit, it is a simple matter to normalize the distribution afterwards. This will be illustrated in an example below.

One last detail to be discussed before giving a detailed example is the problem of choosing a random value from a given Beta distribution.

Sampling from a beta distribution

Let B be a random variable which follows a Beta distribution with parameters α and β. We assume that α and β are both integer values as discussed earlier. We also assume the availability of a routine that generates independent random values from the Uniform (0,1) distribution as well as a routine that generates independent random values from the Standard Normal (0,1) distribution. Both of these routines are readily available on most computer software packages or can be downloaded from the Internet.

First we note that the sum of X independent values from the standard Normal (0,1) distribution follows the Chi-square distribution with K degrees of freedom. Accordingly, we create X_α and X_β as chi-square distributed random variables with 2α and 2β degrees of freedom respectively by sampling 2α and 2β times respectively from the standard Normal (0,1) distribution. We square each of the standard Normal values and add them to produce X_α and X_β.

We then use the following to create the value B which is a randomly sampled value from a Beta distribution with parameters α and β.

$$P = \frac{X_\alpha}{X_\alpha + X_\beta} \tag{B.11}$$

B.3 A numerical example

Let K be the number of sensors. In this example we take $K = 3$. Let N be the (unknown) number of targets in the region/time of interest. Let the observed data be as follows.

Let S_1 be the number of targets observed by sensor 1. In this example we take $S_1 = 3$. Let S_2 be the number of targets observed by sensor 2. In this example we take $S_2 = 7$. Let S_3 be the number of targets observed by sensor 3. In this example we take $S_3 = 10$.

We assume that there is target identification for all sensors and that 13 unique targets are detected. The minimum value of N is thus 13 in the analysis below ($N_{min} = 13$).o

Analysis using non-informative prior distributions

We first consider the analysis when all prior distributions are non-informative.

We assume that the number of targets detected by sensor i, given that there are $N = n$ targets in the region and time period of interest, follows a Binomial distribution with parameter P_i. The unknown parameter P_i is considered to be a random variable in a complete Bayesian analysis. The non-informative prior for P_i is a uniform distribution on (0,1), illustrated in Figure B.1.

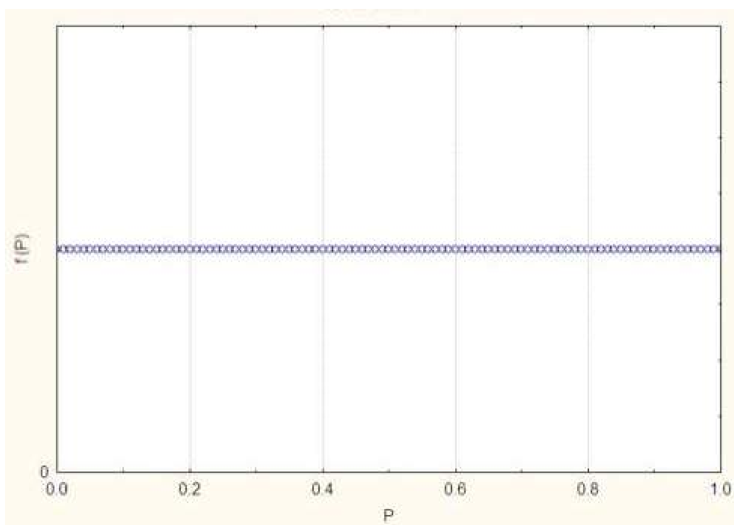

Figure B.1: *Uniform non-informative prior distribution for Pi.*

It is also necessary to provide a prior distribution for the unknown number of targets, N. Figure B.2 presents a reasonable non-informative prior for N.

$$\Pr(N = n) = \frac{1}{n} \quad \textit{for } n = N_{\min}, N_{\min} + 1, \ldots \tag{B.12}$$

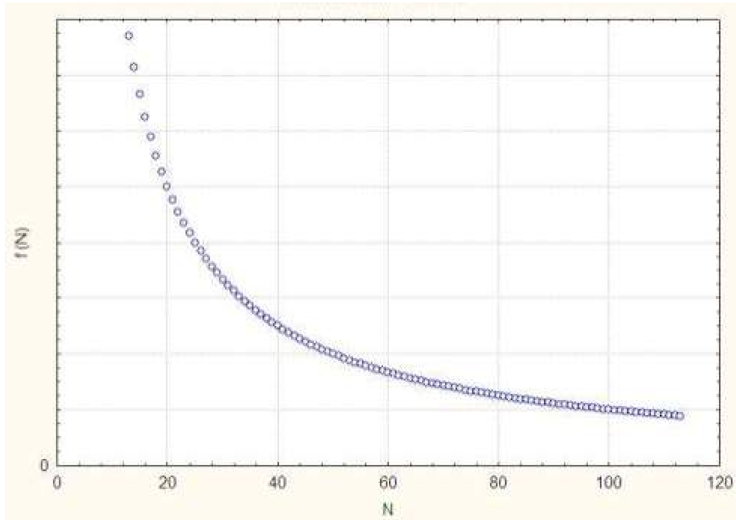

Figure B.2: *Non-informative prior distribution for N.*

We assume conditional independence between the sensors. This means that the number of targets detected by each sensor depends only on how many targets are in the area/time window. Given the number of targets, the number of targets detected by sensor i depends only on the Binomial parameter P_i associated with that sensor.

We can now use (B9) to calculate the posterior relative likelihood for each possible value of N. The multiple integrals are evaluated using Monte Carlo integration. For each of several thousand iterations values of P_1, P_2, P_3, and N are drawn at random from their respective prior distributions and the expression in the equation below is evaluated. Note that this process is self-weighting so that the posterior relative likelihood for each value of N is simply the sum of the values calculated for the several thousand iterations. Note also that the posterior probability distribution for N is derived by normalizing the posterior relative likelihood distribution.

A practical upper limit for N must also be chosen. In this case an upper limit of 60 targets was assumed. As can be seen below the likelihood that N exceeds 60 becomes vanishingly small.

For these selected non-informative priors, the results of the numerical analysis are illustrated in Figure B.3. The number of iterations in the Monte Carlo integration was 10,000.

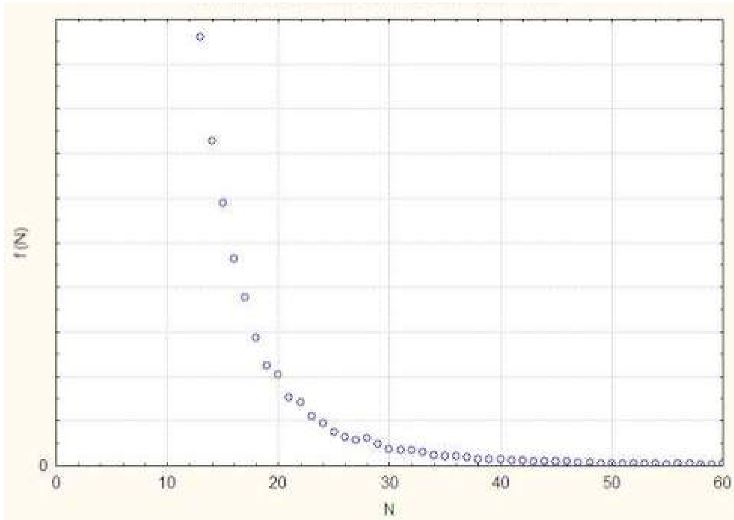

Figure B.3: *Posterior distribution for N with non-informative priors.*

The modal value for *N* is 13. The mean is 18.5 and the median is 15.5. The upper 95% confidence limit for *N* based on the posterior distribution is 33.

Given the nature of the posterior distribution (severely skewed to the right), it is recommended that the median and 95th percentile be chosen as descriptive statistics for comparison purposes. The conclusion would then be: *based only on the current detections the number of targets in the area is estimated to be 15.5 with a minimum of 13 and an upper 95% confidence limit of 33.*

Analysis using mildly-informative prior distributions

In order to assess the effect of prior knowledge on the results, we next perform the analysis with mildly informative prior distributions for the unknown sensor detection probabilities P_1, P_2, and P_3. We continue to use the same non-informative prior distribution for *N*.

We assume that little is known about the detection probability of Sensor 1 except that it may be assumed to be closer to 0.50 than to either 0 or 1. In this case a mildly informative prior is a beta distribution with parameters $\alpha = 2$ and $\beta = 2$. This distribution is shown in Figure B.4. It has a modal value of 0.5 and drops to zero at both 0 and 1.

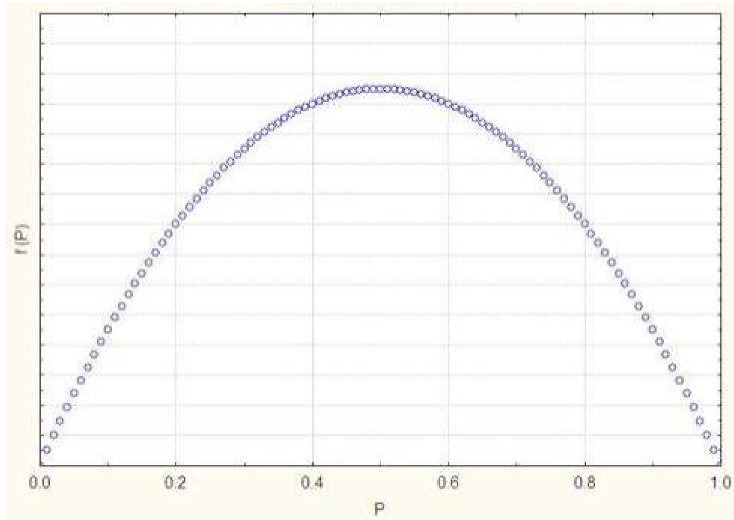

Figure B.4: *Mildly informative prior for P1 (Mode = 0.50).*

For Sensor 2, we suppose that there is evidence from a live trial under conditions sufficiently similar to the current example. In this live trial, Sensor 2 detected and identified two targets out of three. We can incorporate this prior knowledge in the analysis by using a Beta prior for the Sensor 2 detection probability with parameters $\alpha = 3$ and $\beta = 2$. This distribution has a modal value at 0.67. It is illustrated in Figure B.5.

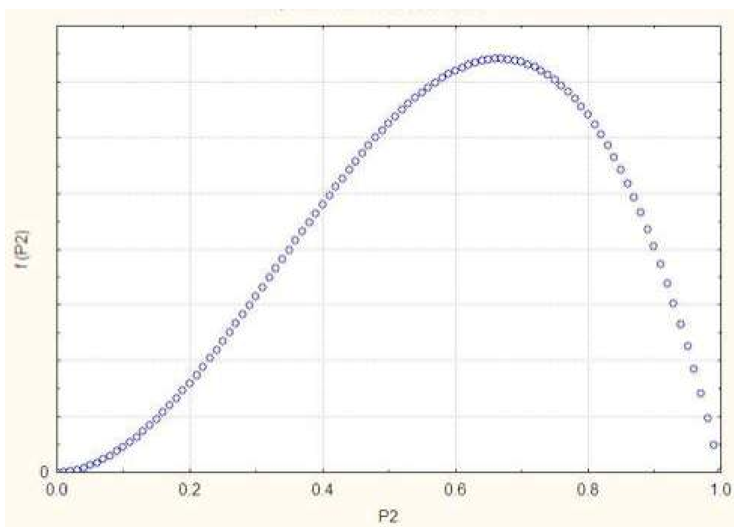

Figure B.5: *Mildly informative prior for P_2 (Mode = 0.67).*

Finally, for Sensor 3 we again suppose that there is evidence from a live trial in which Sensor 3 successfully detected and identified 4 out of 5 known targets. We can incorporate the prior knowledge into the analysis by using a Beta prior for the Sensor 3 detection probability with parameters $\alpha = 5$ and $\beta = 2$, and a modal value at 0.80. It is illustrated in Figure B.6.

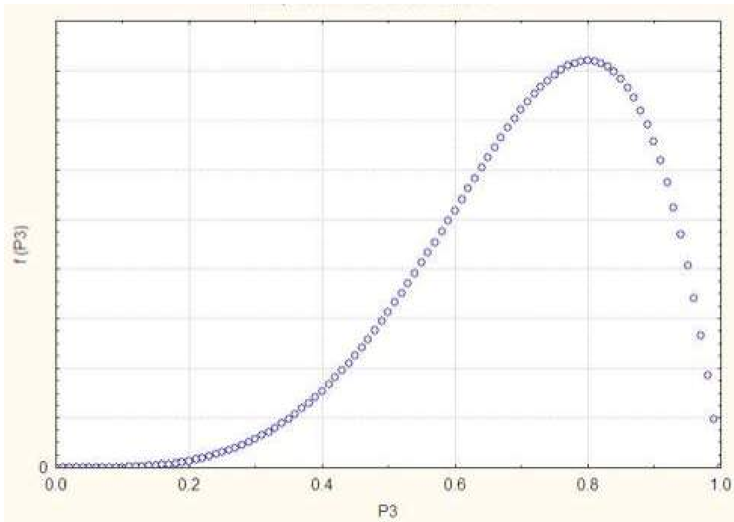

Figure B.6: *Mildly informative prior for P3 (Mode = 0.80).*

Results

Using the same numerical procedure as above, the posterior distribution of *N* was calculated and is shown in Figure B.7.

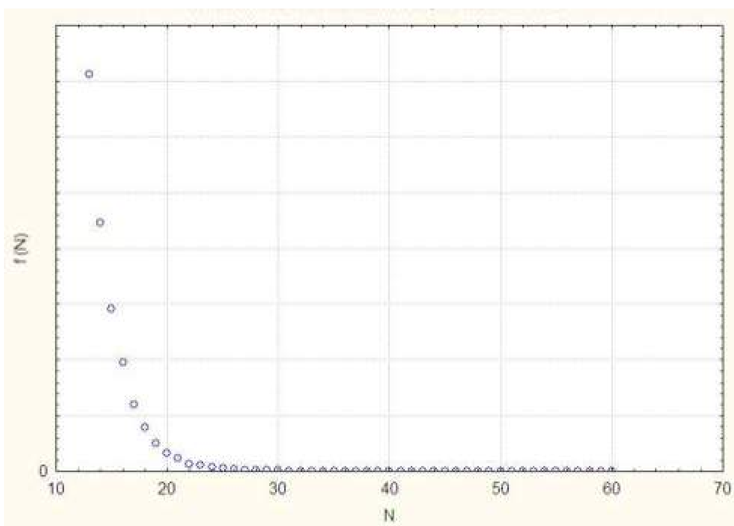

Figure B.7: *Posterior distribution for N with mildly informative priors.*

The effect of even mildly informative prior knowledge on the sensor detection probabilities is evident. The median value is now 13.6 as compared to the previous value of 15.5. Even more dramatic is the reduction in the upper 95% confidence limit to 20 from the previous value of 33.

This page intentionally left blank.

List of symbols/abbreviations/acronyms/initialisms

AIS	Automatic Identification System
AOI	Area of Interest
ARP	Advanced Research Project
CF	Canadian Forces
CORA	Centre for Operational Research and Analysis
DND	Department of National Defence
DRDC	Defence Research & Development Canada
JTFP	Joint Task Force Pacific
LRIT	Long Range Identification & Tracking
MCPG	Maritime Capability Planning Guidance
MDA	Maritime Domain Awareness
NB()	Negative Binomial
RMP	Recognized Maritime Picture
RS2	RADARSAT 2
SA	Situational Awareness
SAW	Surveillance Analysis Workbook
VHF	Very High Frequency

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Centre sponsoring a contractor's report, or tasking agency, are entered in Section 8.) Defence Research and Development Canada – CORA 101 Colonel By Drive Ottawa, Ontario K1A 0K2	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) UNCLASSIFIED 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC DECEMBER 2012

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

A Bayesian approach to estimating detection performance in a multi-sensor environment

4. AUTHORS (last name, followed by initials – ranks, titles, etc., not to be used)

Horn, S.A.; Mason, D.W.; Zegers, A.; Emond, E.J.

5. DATE OF PUBLICATION (Month and year of publication of document.) September 2014	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 56	6b. NO. OF REFS (Total cited in document.) 25

7. DESCRIPTIVE NOTES (The category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g., interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Scientific Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence Research and Development Canada – CORA
101 Colonel By Drive
Ottawa, Ontario K1A 0K2

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC-RDDC-2014-R56	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

This report presents the results of Operational Research support to maritime surveillance operations in the Canadian maritime approaches. The principal development is a Bayesian method to estimate the performance of sensors in a way that can be applied during ongoing operations. This novel method does not require knowledge of the sea-truth to evaluate sensors. A method to estimate sea-truth is also presented, which provides a new and unique analysis capability for the estimation of surveillance gaps. The series of new metrics and methods are implemented in a variety of operational tools for analysis of the recognized maritime picture (RMP), which support new operational procedures and processes. The desired result is a more robust and optimized maritime surveillance capability.

----

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g., Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Maritime Domain Awareness, Surveillance, Bayesian Statistics, Multi-Sensor